



The Efficacy of the Government's Use of Past Performance Information: An Exploratory Study

11th Annual Acquisition Research Symposium

Tim Hawkins, Lt Col, USAF (Ret.), Ph.D.

15 May 2014



Issues in Past Performance Evaluation

Purpose & Research Question

Theoretical Frameworks

- Agency Theory
- Organizational Behavior
- Channel Communication

Methodology

Results

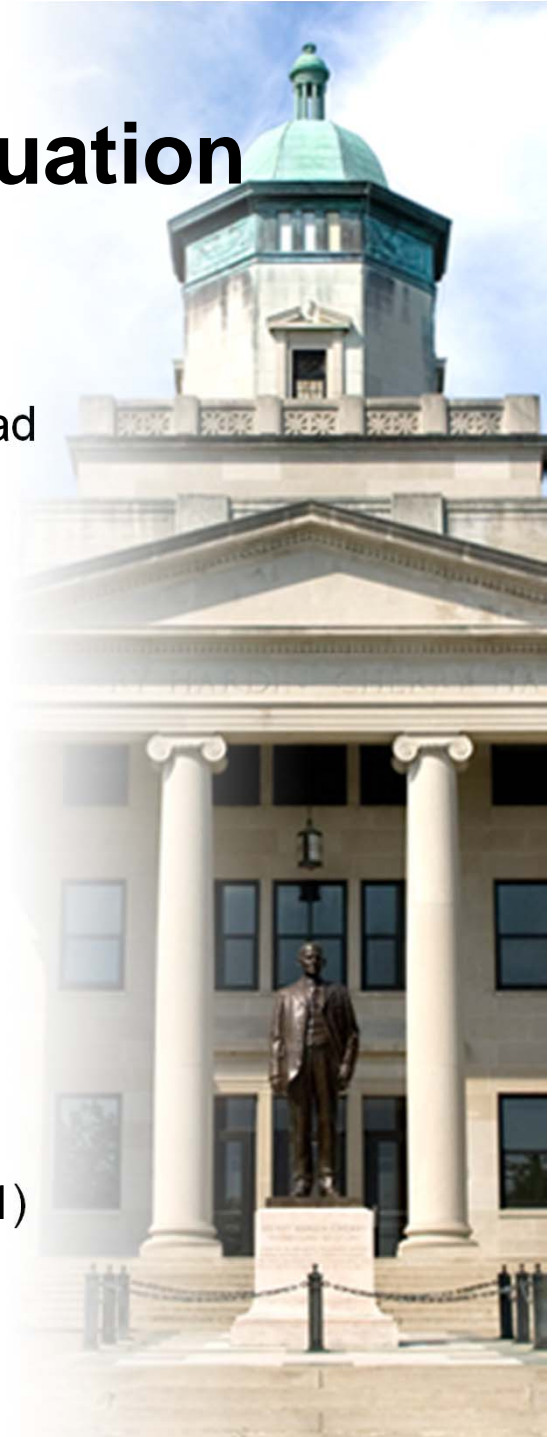
Implications





Issues in Past Performance Evaluation

- Only 31% of contract actions requiring CPARS reporting had completed reports (GAO,2009)
- Insufficient information to support ratings (OFPP, 2011)
 - how the contractor met, exceeded, or failed to meet requirements
- Incomplete reports – some categories not rated
- “*Halo Effect*” - Raters often inflate ratings to avoid conflict with the contractor (GAO, 2009)
 - PP increasingly subject to Contract Disputes Act
- Much attention and some improvement recently
 - Fed Gov’t PP Guide (2012), formerly DoD Guide (2011)



Problems

- Degree of inaccuracy of PPI unknown
 - Inaccurate PP assessments can harm contractors' reputations
 - Can bias source selections resulting in *adverse selection*.
- Reasons for inaccuracy not empirically explored
- Transaction costs not insignificant – but unknown precisely
 - If PPI is not reliable, and if evaluators cannot use the PPI to discriminate between proposals (Kelman, 2010), the effort of collecting and reporting, then later evaluating and documenting PPI is squandered
 - Federal contract managers are overworked (GAO, 2009) and understaffed (GAO, 2001)
 - Awarded 5.9M contract actions at \$538B in FY10
- PP evals, thus, often add little value to selection decisions



Purpose & Research Question

Purpose: Explore the efficacy of the government's current use of PPI

- Validate reported issues
- Tee up future research

Research Questions:

- Are PP reports useful?
 - Motivate suppliers to perform?
 - Reduce future performance uncertainty?
 - To what extent do PP evaluations/ratings influence source selection decisions?
- Why do PP evaluations/ratings lack sufficient justification?
- Why are PP evaluations sometimes inaccurate?
- In the cases of multiple evaluators on a single contract action, do PP evaluations/ratings deviate among evaluators, and, if so, why?
- Why do reviewing officials change the ratings of the evaluator (assessing official)?



Agency Theory – 2 problems:

1. conflicting interests between principal and agent and
 2. difficulty and cost associated with monitoring agents, and the associated uncertainty for not having perfect information (Eisenhardt, 1989).
- Supplier as agent to buyer
 - Evaluator and other stakeholders as agents to buyer
 - Allegiance to buyer, program, or ktr (fairness; effect on ktr)?

Organizational Behavior

- PP likened to employee evaluations
 - Multiple raters
 - Halo effect

Channel Communication

- Formal Comm decr distortion (Mohr & Sohi, 1995)



Qualitative – appropriate when:

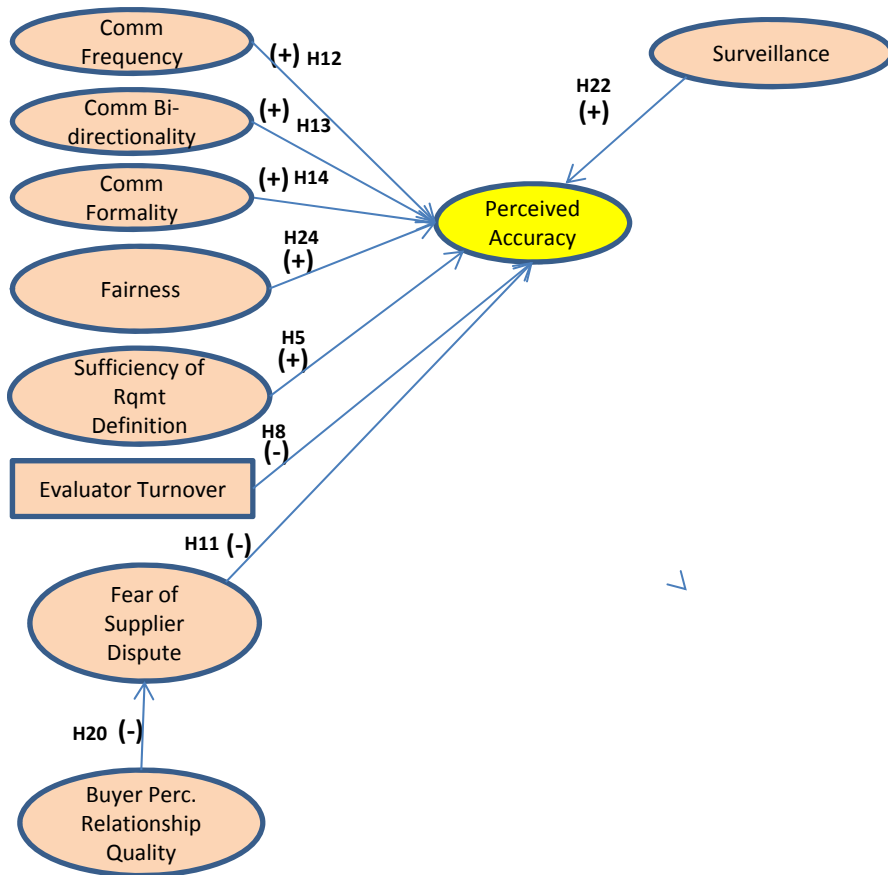
1. research is exploratory in nature (“why?”)
2. researcher has no control of the behavioral events being researched
3. focus is on contemporary events

Data Collection

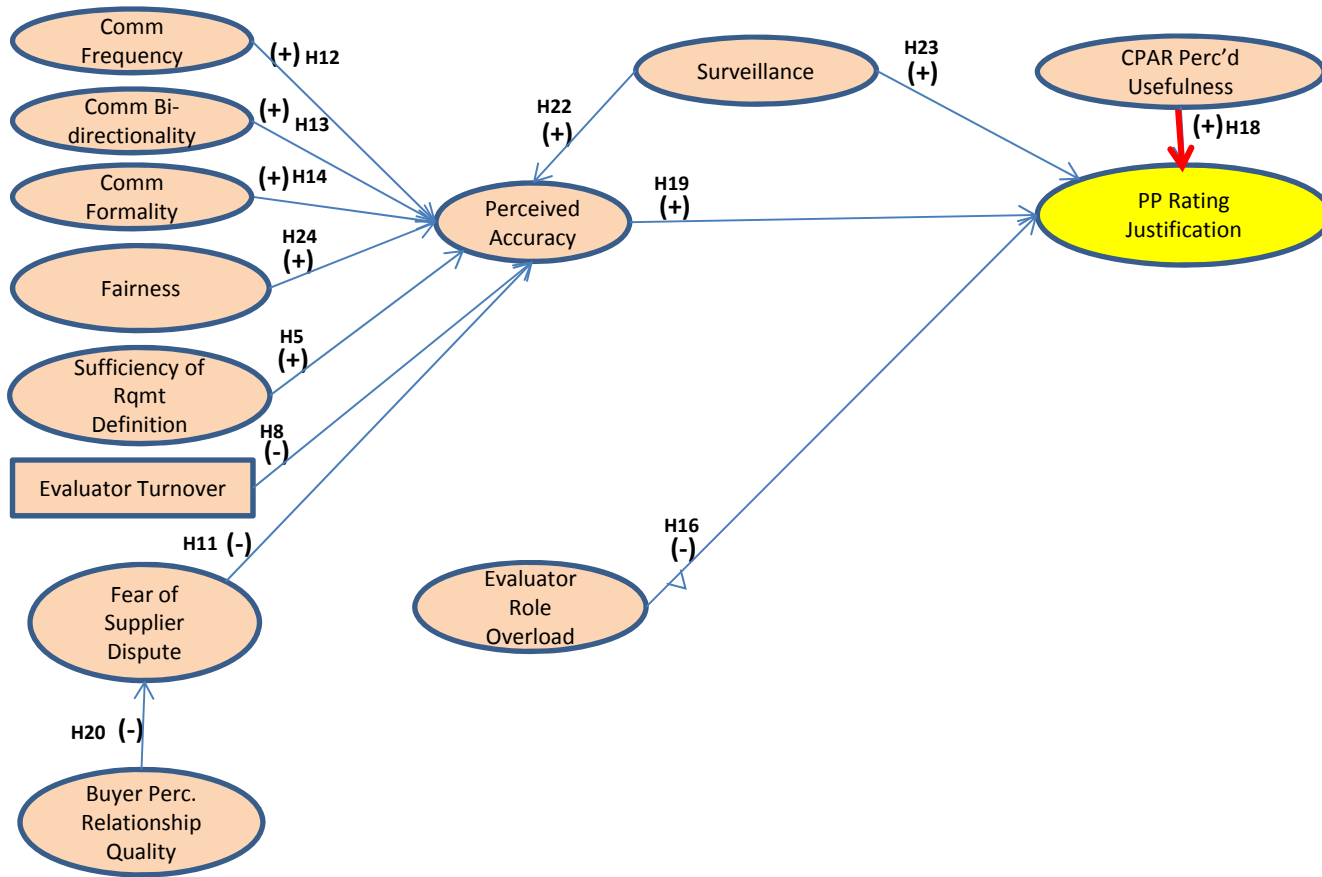
- Interview Protocol
- 8 Interviews
- 38-67 Min; avg 18 pages transcribed



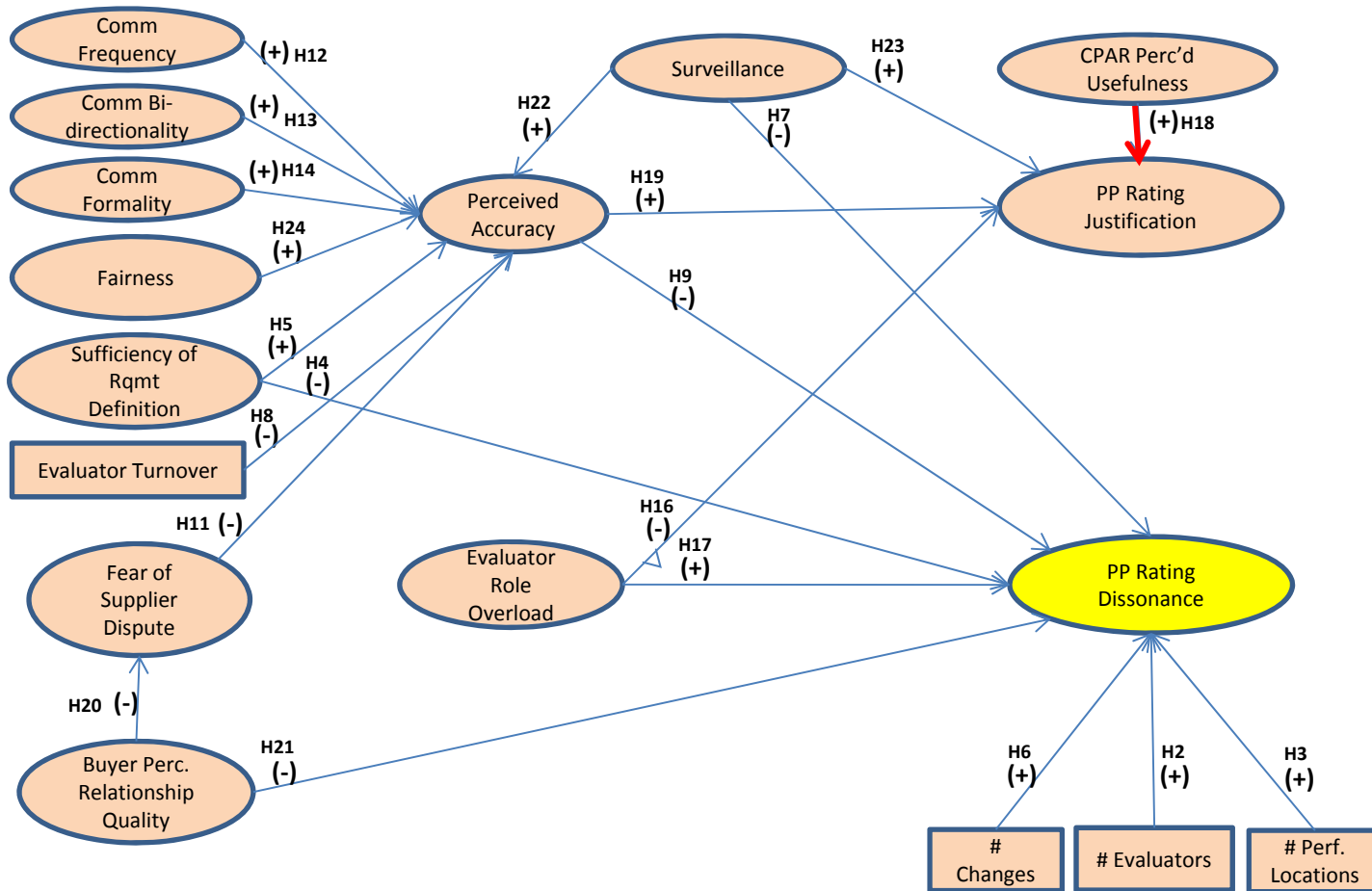
Results – Conceptual Model



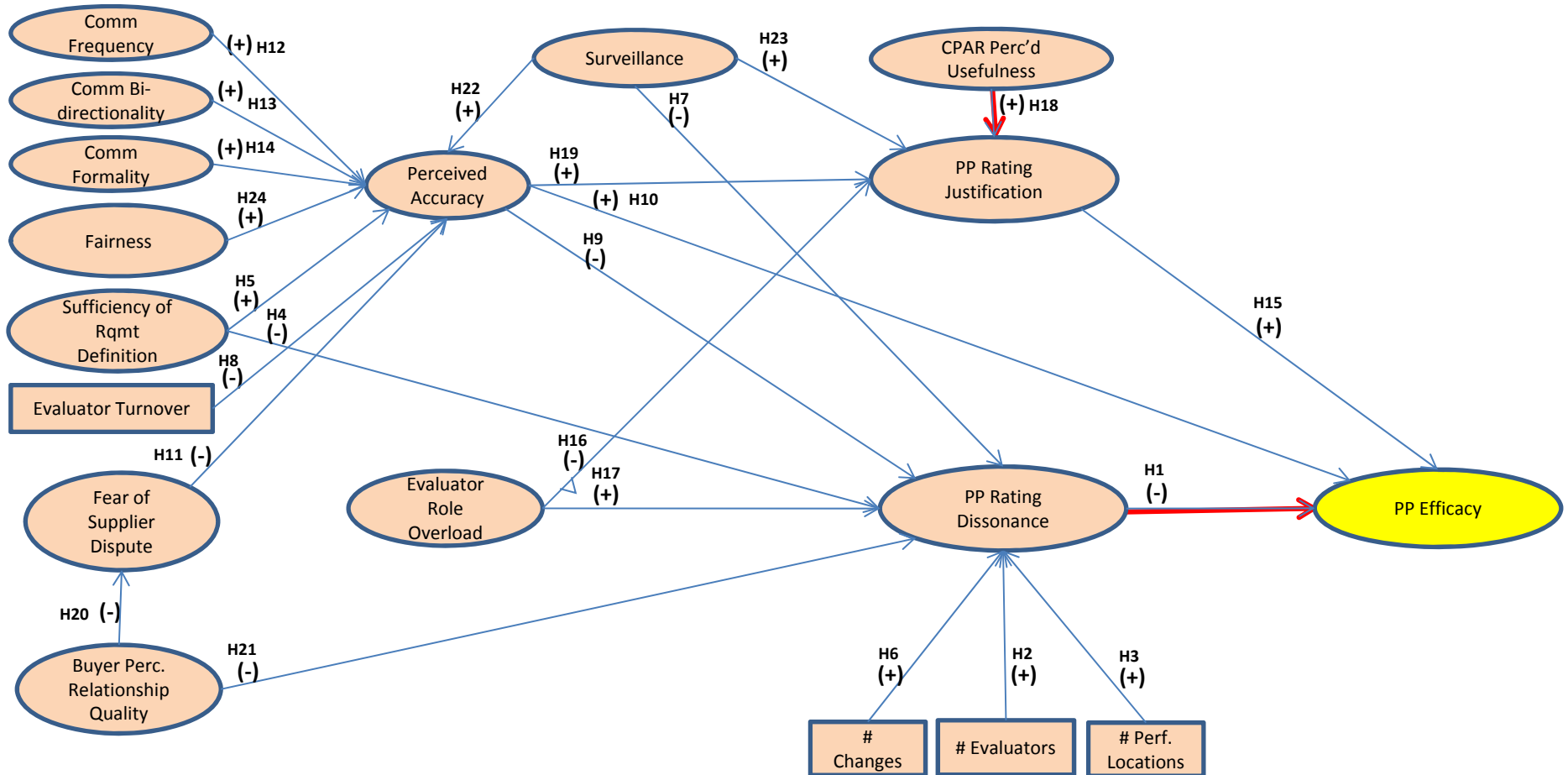
Results – Conceptual Model



Results – Conceptual Model



Results – Conceptual Model



Implications

- On average, past performance evaluations consume nearly 38 man-hrs of effort (rng 8-100).
- Leaders should ensure evaluators have sufficient time to perform their PP evals
 - Manning models need to account for PP workload
- Thoroughly define requirements—including measurements of success and precise definitions of CPAR ratings tailored to the requirement—prior to solicitation
- PP reporting is often not a sufficient surrogate for contractor performance management.
 - More frequent, formal, and two-way communication appears necessary to ensure rating accuracy



Implications (cont.)

- Leaders should mitigate turnover of performance evaluators
- Independent agency *requirements* ombudsman to ensure sufficient definitions of rqmts and PP ratings
- Halo effect confirmed – due to fear of dispute, fairness, protecting program, and concern for effect on ktr – particularly whether one instance of perf failure should represent all other successful perf opportunities
- Some use PP as leverage – as a threat ex ante, or as punishment ex post
- Lots of variance in performance info collection, recording, and sharing
 - Consider SPE system



Conclusion

- Confirmed many reported weaknesses
- Explained why the systemic weaknesses occur:
 - Accuracy of performance info,
 - Workload,
 - Variance in communications,
 - Poor rating justifications,
 - Variance in performance info collection, reporting, and sharing
- But, need to quantitatively confirm findings with a large sample
- Future research could explore effects of low PP efficacy on the contractor:
 - Performance?
 - Relationship quality?

